

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:32:59 ; Search time 46 Seconds  
(without alignments)  
108.493 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1279676 seqs, 311918243 residues  
Total number of hits satisfying chosen parameters: 1279676

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	16	14	US-10-001-938-3
2	86	100.0	340	9	US-09-764-868-1153
3	86	100.0	341	9	US-09-764-868-1152
4	86	100.0	376	15	US-10-369-493-689
5	86	100.0	378	15	US-10-369-493-386
6	86	100.0	378	15	US-10-369-493-21305
7	81	94.2	15	9	US-09-756-983-10
8	81	94.2	15	9	US-09-828-574-10
9	81	94.2	15	14	US-10-001-938-27
10	81	94.2	15	14	US-10-299-540-4
11	81	94.2	15	14	US-10-299-184-4
12	81	94.2	15	14	US-10-239-313A-444
13	81	94.2	16	14	US-10-239-313A-447
14	79	91.9	365	15	US-10-369-493-10648
15	78	90.7	376	15	US-10-369-493-12087

16	72	83.7	15	9	US-09-828-574-11	Sequence 11, Appl
17	72	83.7	15	14	US-10-299-540-5	Sequence 5, Appl
18	72	83.7	15	14	US-10-299-184-5	Sequence 5, Appl
19	69	80.2	382	15	US-10-369-493-8065	Sequence 8065, Ap
20	69	80.2	384	15	US-10-369-493-21849	Sequence 21849, A
21	67	77.9	379	15	US-10-369-493-7330	Sequence 7330, Ap
22	67	77.9	380	15	US-10-369-493-4573	Sequence 4573, Ap
23	63	73.3	376	15	US-10-369-493-47	Sequence 47, Appl
24	62	72.1	374	15	US-10-369-493-13658	Sequence 13658, A
25	61	70.9	378	15	US-10-369-493-8450	Sequence 8450, Ap
26	60	69.8	368	9	US-09-861-451A-2	Sequence 2, Appl
27	60	69.8	375	15	US-10-369-493-15859	Sequence 15859, A
28	60	69.8	376	15	US-10-369-493-15487	Sequence 15487, A
29	59	68.6	358	15	US-10-369-493-9641	Sequence 9641, Ap
30	58	67.4	379	15	US-10-369-493-18523	Sequence 18523, A
31	58	67.4	385	15	US-10-369-493-16754	Sequence 16754, A
32	57	66.3	369	15	US-10-369-493-19646	Sequence 19646, A
33	56	65.1	362	15	US-10-369-493-9305	Sequence 9305, Ap
34	56	65.1	378	16	US-10-474-776-634	Sequence 634, App
35	54	62.8	15	14	US-10-299-540-6	Sequence 6, Appl
36	54	62.8	15	14	US-10-299-184-6	Sequence 6, Appl
37	53	61.6	423	15	US-10-369-493-12619	Sequence 12619, A
38	52	60.5	360	15	US-10-369-493-10371	Sequence 10371, A
39	52	60.5	372	15	US-10-369-493-23206	Sequence 23206, A
40	52	60.5	376	15	US-10-369-493-1212	Sequence 1212, Ap
41	50	58.1	297	12	US-10-282-122A-54551	Sequence 54551, A
42	50	58.1	362	16	US-10-437-963-190831	Sequence 190831, A
43	50	58.1	383	16	US-10-437-963-108162	Sequence 108162, A
44	50	58.1	404	12	US-10-425-114-59964	Sequence 59964, A
45	49	57.0	146	12	US-10-424-599-191916	Sequence 191916, A

ALIGNMENTS

RESULT 1  
US-10-001-938-3  
; Sequence 3, Application US/10001938  
; Publication No. US20030031679A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: ALBANI, Salvatore  
; APPLICANT: CARSON, Dennis  
; APPLICANT: PRAKEN, Berent  
; APPLICANT: MARTINI, Alberto  
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES DERIVED FROM HEAT SHOCK PROTEINS AND USI  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: UCSD1360-1  
; CURRENT APPLICATION NUMBER: US/10/001, 938  
; PRIOR FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US 60/245,181  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-001-938-3

Query Match 100.0%; Score 86; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 9.3e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAYDQYGHAAFEQ 16  
DB 1 QKRAYDQYGHAAFEQ 16

RESULT 2  
US-09-764-868-1153  
; Sequence 1153, Application US/09764868  
; Patent No. US20020168711A1

```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1153
; LENGTH: 340
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (110)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (156)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (196)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (198)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (216)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-09-764-868-1153

Query Match          100.0%; Score 86; DB 9; Length 340;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16
  |||||
Db 25 QKRAAYDQYGHAAFEQ 40

RESULT 3
US-09-764-868-1152
; Sequence 1152, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1152
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-764-868-1152

Query Match          100.0%; Score 86; DB 9; Length 341;
Best Local Similarity 100.0%; Pred. No. 2.2e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16
  |||||
Db 26 QKRAAYDQYGHAAFEQ 41

RESULT 4
US-10-369-493-689
```

```
; Sequence 689, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 689
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-10-369-493-689

Query Match          100.0%; Score 86; DB 15; Length 376;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16
  |||||
Db 61 QKRAAYDQYGHAAFEQ 76

RESULT 5
US-10-369-493-386
; Sequence 386, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 386
; LENGTH: 378
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
; US-10-369-493-386

Query Match          100.0%; Score 86; DB 15; Length 378;
Best Local Similarity 100.0%; Pred. No. 2.4e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16
  |||||
Db 61 QKRAAYDQYGHAAFEQ 76

RESULT 6
US-10-369-493-21305
; Sequence 21305, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; US-10-369-493-21305
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:30:33 ; Search time 23 Seconds  
(without alignments)  
35.914 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAAYDOYGHAFFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	131	4	US-09-553-498-4
2	86	100.0	131	4	US-09-618-869-4
3	86	100.0	380	4	US-09-543-681A-5912
4	86	100.0	380	4	US-09-489-039A-9011
5	86	100.0	399	4	US-09-553-498-2
6	86	100.0	399	4	US-09-618-869-2
7	81	94.2	15	1	US-08-618-464-4
8	81	94.2	15	3	US-09-107-615-4
9	81	94.2	15	5	PCT-US95-04896-4
10	72	83.7	15	1	US-08-618-464-5
11	72	83.7	15	3	US-09-107-615-5
12	72	83.7	15	5	PCT-US95-04896-5
13	62	72.1	381	4	US-09-252-991A-27174
14	61	70.9	407	4	US-09-540-236-2899
15	56	65.1	352	2	US-08-472-534-6
16	54	62.8	15	1	US-08-618-464-6
17	54	62.8	15	3	US-09-107-615-6
18	52	60.5	375	4	US-09-328-352-4984
19	47	54.7	385	4	US-09-134-001C-3688
20	46	53.5	419	2	US-08-686-417-3
21	45	52.3	677	4	US-09-252-991A-18102
22	44	51.2	915	4	US-09-252-991A-24992
23	43	50.0	1895	2	US-08-619-554-4
24	41	47.7	87	4	US-09-882-835-4
25	41	47.7	223	4	US-09-658-644-4
26	41	47.7	348	2	US-08-974-546-1
27	41	47.7	387	4	US-09-543-681A-7130

28	41	47.7	845	4	US-09-252-991A-29230	Sequence 29230, A
29	40	46.5	317	4	US-09-328-352-8169	Sequence 8169, Ap
30	40	46.5	487	4	US-09-252-991A-21980	Sequence 21980, A
31	38	44.2	107	1	US-08-326-117B-3	Sequence 3, Appli
32	38	44.2	107	3	US-08-982-129-3	Sequence 3, Appli
33	38	44.2	108	3	US-08-893-534A-5	Sequence 5, Appli
34	38	44.2	108	3	US-08-996-679-5	Sequence 5, Appli
35	38	44.2	108	3	US-08-939-853A-7	Sequence 7, Appli
36	38	44.2	108	3	US-09-115-395-5	Sequence 5, Appli
37	38	44.2	108	3	US-09-113-977C-38	Sequence 38, Appli
38	38	44.2	108	4	US-09-507-102-5	Sequence 5, Appli
39	38	44.2	108	4	US-09-250-059-5	Sequence 5, Appli
40	38	44.2	108	4	US-09-248-074-5	Sequence 5, Appli
41	38	44.2	108	4	US-09-357-717-5	Sequence 5, Appli
42	38	44.2	108	4	US-09-458-870-5	Sequence 5, Appli
43	38	44.2	108	4	US-09-351-048A-38	Sequence 38, Appli
44	38	44.2	108	4	US-09-248-015-5	Sequence 5, Appli
45	38	44.2	108	4	US-09-544-782-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1  
US-09-553-498-4  
; Sequence 4, Application US/09553498  
; Patent No. 6309861  
; GENERAL INFORMATION:  
; APPLICANT: Ambrosius, Dorothee  
; APPLICANT: Rudolph, Rainer  
; APPLICANT: Schaeffner, Joerg  
; APPLICANT: Schwarz, Elisabeth  
; TITLE OF INVENTION: Process for the production of naturally folded and secreted prote:  
; FILE REFERENCE: Case 20379  
; CURRENT APPLICATION NUMBER: US/09/553,498  
; CURRENT FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: EP99107412.1  
; PRIOR FILING DATE: 1999-04-26  
; NUMBER OF SEQ ID NOS: 10  
; SEQ ID NO 4  
; LENGTH: 131  
; TYPE: PRT  
; ORGANISM: E. coli  
US-09-553-498-4

Query Match 100.0%; Score 86; DB 4; Length 131;  
Best Local Similarity 100.0%; Pred. No. 7.8e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDOYGHAFFEQ 16  
Db 84 QKRAAYDOYGHAFFEQ 99

RESULT 2  
US-09-618-869-4  
; Sequence 4, Application US/09618869  
; Patent No. 6455279  
; GENERAL INFORMATION:  
; APPLICANT: Ambrosius, Dorothee  
; APPLICANT: Rudolph, Rainer  
; APPLICANT: Schaeffner, Joerg  
; APPLICANT: Schwarz, Elisabeth  
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND  
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR  
; FILE REFERENCE: 20381  
; CURRENT APPLICATION NUMBER: US/09/618,869  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: EP99114811.5  
; PRIOR FILING DATE: 1999-07-29  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn Ver. 2.1

```
; SEQ ID NO 4
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-618-869-4
```

```
Query Match          100.0%; Score 86; DB 4; Length 131;
Best Local Similarity 100.0%; Pred. No. 7.8e-08;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QKRAAYDQYGHAAFEQ 16
         |||||
Db       84 QKRAAYDQYGHAAFEQ 99
```

## RESULT 3

```
US-09-543-681A-5912
; Sequence 5912, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5912
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-5912
```

```
Query Match          100.0%; Score 86; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QKRAAYDQYGHAAFEQ 16
         |||||
Db       63 QKRAAYDQYGHAAFEQ 78
```

## RESULT 4

```
US-09-489-039A-9011
; Sequence 9011, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9011
; LENGTH: 380
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9011
```

```
Query Match          100.0%; Score 86; DB 4; Length 380;
Best Local Similarity 100.0%; Pred. No. 2.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QKRAAYDQYGHAAFEQ 16
         |||||
Db       64 QKRAAYDQYGHAAFEQ 79
```

```
RESULT 5
US-09-553-498-2
; Sequence 2, Application US/09553498
; Patent No. 6309861
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: Process for the production of naturally folded and secreted prote
```

```
; FILE REFERENCE: Case 20379
; CURRENT APPLICATION NUMBER: US/09/553,498
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: EP99107412.1
; PRIOR FILING DATE: 1999-04-26
; NUMBER OF SEQ ID NOS: 10
; SEQ ID NO 2
; LENGTH: 399
; TYPE: PRT
; ORGANISM: E. coli
US-09-553-498-2
```

```
Query Match          100.0%; Score 86; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QKRAAYDQYGHAAFEQ 16
         |||||
Db       84 QKRAAYDQYGHAAFEQ 99
```

## RESULT 6

```
US-09-618-869-2
; Sequence 2, Application US/09618869
; Patent No. 6455279
; GENERAL INFORMATION:
; APPLICANT: Ambrosius, Dorothee
; APPLICANT: Rudolph, Rainer
; APPLICANT: Schaeffner, Joerg
; APPLICANT: Schwarz, Elisabeth
; TITLE OF INVENTION: PROCESS FOR THE PRODUCTION OF NATURALLY FOLDED AND
; TITLE OF INVENTION: SECRETED PROTEINS BY CO-SECRETION OF MOLECULAR
; FILE REFERENCE: 20381
; CURRENT APPLICATION NUMBER: US/09/618,869
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: EP99114811.5
; PRIOR FILING DATE: 1999-07-29
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-618-869-2
```

```
Query Match          100.0%; Score 86; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 QKRAAYDQYGHAAFEQ 16
         |||||
Db       84 QKRAAYDQYGHAAFEQ 99
```

## RESULT 7

```
US-08-618-464-4
; Sequence 4, Application US/08618464
; Patent No. 573570
; GENERAL INFORMATION:
; APPLICANT: CARSON, DENNIS A.
; APPLICANT: ALBANI, SALVATORE
; TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:25:58 ; Search time 11 Seconds  
(without alignments)  
75.738 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAYDQYGHAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	375	1 DNAS_ECOLI	P08622 escherichia
2	86	100.0	377	1 DNAS_HAEDU	P48208 haemophilus
3	86	100.0	378	1 DNAS_SALTU	Q60004 salmonella
4	86	100.0	382	1 DNAS_HAEIN	P43735 haemophilus
5	83	96.5	375	1 DNAS_ACTAC	P77866 actinobacil
6	83	96.5	381	1 DNAS_VIBCH	O34242 vibrio chol
7	83	96.5	385	1 DNAS_VIBHA	O87385 vibrio harv
8	82	95.3	392	1 NOLC_RHIFR	P26508 rhizobium f
9	80	93.0	372	1 DNAS_PASMT	O9cms2 pasteurella
10	77	89.5	234	1 DNAS_RHILE	O33529 rhizobium 1
11	77	89.5	379	1 DNAS_RHOSP	O08356 rhodopseudo
12	75	87.2	377	1 DNAS_BUCAT	O32465 buchnera ap
13	75	87.2	379	1 DNAS_PASHA	O52065 pasteurella
14	75	87.2	383	1 DNAS_BUCBP	O89au7 buchnera ap
15	74	86.0	378	1 DNAS_BUCAP	Q8K9Y9 buchnera ap
16	73	84.9	370	1 DNAS_ERYRH	Q05646 erysipeloch
17	73	84.9	373	1 DNAS_NEIMA	P57107 neisseria m
18	73	84.9	375	1 DNAS_BRUOV	Q05980 brucella oy
19	73	84.9	377	1 DNAS_AGRIS	P50018 agrobacteri
20	73	84.9	377	1 DNAS_BRAJA	P94319 bradyrhizob
21	73	84.9	377	1 DNAS_BRUME	O8ye77 brucella me
22	73	84.9	377	1 DNAS_BRUSU	O8fxk1 brucella su
23	69	80.2	384	1 DNAS_RHOCQ	O52702 rhodobacter
24	67	77.9	371	1 DNAS_METBS	Q9zfc5 methyllovoru
25	66	76.7	383	1 DNAS_LACSK	O87778 lactobacill
26	64	74.4	370	1 DNAS_RICPR	Q9zdy0 rickettsia
27	64	74.4	374	1 DNAS_COXBU	P42381 coxiella bu
28	63	73.3	376	1 DNAS_AQUAE	O66921 aquifex aeo
29	63	73.3	379	1 DNAS_LEGPN	P50025 legionella
30	62	72.1	377	1 DNAS_PSEAE	Q9hV44 pseudomonas
31	58	67.4	379	1 DNAS_LACLA	P35514 lactococcus
32	58	67.4	385	1 DNAS_CAUCR	P22305 caulobacter
33	57	66.3	364	1 DNAS_BORBU	P28616 borrelia bu

34	57	66.3	369	1 DNAS_NITEU	O06431 nitrosomona
35	57	66.3	376	1 DNAS_LISIN	Q92bn9 listeria in
36	57	66.3	377	1 DNAS_LISMO	Q9s5a3 listeria mo
37	56	65.1	368	1 DNAS_XYLFA	Q9pb06 xylella fas
38	56	65.1	368	1 DNAS_XYLFT	Q87bs9 xylella fas
39	56	65.1	378	1 DNAS_STRPN	P55830 streptococc
40	54	62.8	383	1 DNAS_PORGI	Q9xcac porphyromon
41	53	61.6	307	1 DNAS_DEIFR	O34136 deिनococcus
42	53	61.6	371	1 DNAS_FRATU	P48207 francisella
43	52	60.5	372	1 DNAS_BACSU	P17631 bacillus su
44	52	60.5	376	1 DNAS_METTH	O27352 methanobact
45	52	60.5	384	1 DNAS_HALME	Q9hnb8 halobacteri

ALIGNMENTS

RESULT 1  
DNAS\_ECOLI ID DNAS\_ECOLI STANDARD; PRT; 375 AA.  
AC P08622;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Chapterone protein dnaJ (Heat shock protein J) (HSP40).  
GN DNAS OR GRGP OR B0015.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.  
RX MEDLINE=86111849; PubMed=3003084;  
RA Onki M., Tamura F., Nishimura S., Uchida H.;  
RT "Nucleotide sequence of the Escherichia coli dnaJ gene and  
RT purification of the gene product."  
RL J. Biol. Chem. 261:1778-1781(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=86111850; PubMed=3003085;  
RA Bardwell J.C.A., Tilly K., Craig E., King J., Zyllicz M.,  
RT Georgopoulos C.;  
RT "The nucleotide sequence of the Escherichia coli K12 dnaJ+ gene. A  
RT gene that encodes a heat shock protein."  
RL J. Biol. Chem. 261:1782-1785(1986).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RX MEDLINE=92334977; PubMed=1630901;  
RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
RA Isono K., Mizobuchi K., Nakata A.;  
RT "Systematic sequencing of the Escherichia coli genome: analysis of  
RT the 0-2.4 min region."  
RL Nucleic Acids Res. 20:3305-3308(1992).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Coliádo-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1453-1474(1997).  
RN [5]  
RP POSSIBLE FUNCTION.  
RX MEDLINE=91187894; PubMed=1826368;  
RA Liberek K., Marszalek J., Ang D., Georgopoulos C., Zyllicz M.;  
RT "Escherichia coli DnaJ and GrpE heat shock proteins jointly stimulate  
RT ATPase activity of DnaK."  
RL Proc. Natl. Acad. Sci. U.S.A. 88:2874-2878(1991).  
RN [6]

RP STRUCTURE BY NMR OF 1-107.  
RX MEDLINE=96291434; PubMed=8764403;  
RA Pellechia M., Szyperski T., Wall D., Georgopoulos C., Wuehrich K.;  
RT "NMR structure of the J-domain and the Gly/Phe-rich region of the  
RL Escherichia coli DnaJ chaperone.";  
J. Mol. Biol. 260:236-250(1996).  
RN [7]  
RP STRUCTURE BY NMR OF 1-104.  
RX MEDLINE=99224904; PubMed=10210198;  
RA Huang K., Planagan J.M., Prestegard J.H.;  
RT "The influence of C-terminal extension on the structure of the 'J-  
RL domain' in E. coli DnaJ.";  
Protein Sci. 8:203-214(1999).  
RN [8]  
RP STRUCTURE BY NMR OF 130-208.  
RX MEDLINE=20351465; PubMed=10891270;  
RA Martinez-Yamout M., Legge G.B., Zhang O., Wright P.E., Dyson H.J.;  
RT "Solution structure of the cysteine-rich domain of the Escherichia  
RL coli chaperone protein DnaJ.";  
J. Mol. Biol. 300:805-818(2000).  
CC -!- FUNCTION: INTERACTS WITH DNAK, TO DISASSEMBLE A PROTEIN COMPLEX  
AT THE PHAGE LAMBDA ORIGIN OF REPLICATION. STIMULATES, JOINTLY  
CC WITH GRPE, THE ATPASE ACTIVITY OF DNAK.  
CC -!- COFACTOR: Binds 2 zinc ions per monomer.  
CC -!- SUBUNIT: Homodimer.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- INDUCTION: BY HEAT SHOCK UNDER THE CONTROL OF THE HTPR REGULATORY  
CC PROTEIN.  
CC -!- SIMILARITY: Belongs to the dnaJ family.  
CC -!- SIMILARITY: Contains 1 J domain.  
CC -!- SIMILARITY: Contains 1 CR domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M12565; AAA23693.1; -.  
DR EMBL; D10483; BAB96590.1; -.  
DR EMBL; AE000112; AAC73126.1; -.  
DR PIR; A92572; HHECDJ.  
DR PDB; 1XBL; 11-JAN-97.  
DR PDB; 1BQ0; 15-JUN-99.  
DR PDB; 1BQ2; 15-JUN-99.  
DR PDB; 1EXK; 26-JUL-00.  
DR ECODBASE; H036.5; 6TH EDITION.  
DR EcoGene; EG10240; dnaJ.  
DR InterPro; IPR002939; DnaJ\_C.  
DR InterPro; IPR001305; DnaJ\_CXXCXGXG.  
DR InterPro; IPR001623; DnaJ\_N.  
DR InterPro; IPR008971; HSP40\_DnaJ\_pep.  
DR InterPro; IPR003095; Hsp\_DnaJ.  
DR Pfam; PF00226; DnaJ; 1.  
DR Pfam; PF01556; DnaJ\_C; 1.  
DR Pfam; PF00684; DnaJ\_CXXCXGXG; 1.  
DR PRINTS; PR00625; DNAJPROTEIN.  
DR SMART; SM00271; DnaJ; 1.  
DR PROSITE; PS00636; DNAJ\_1; 1.  
DR PROSITE; PS00076; DNAJ\_2; 1.  
DR PROSITE; PS00637; DNAJ\_CXXCXGXG; 1.  
KW Chaperone; DNA replication; Heat shock; Repeat; Zinc; Metal-binding;  
KW 3D-structure; Complete proteome.  
FT INIT MET 0  
FT DOMAIN 2  
FT DOMAIN 76 71 J-DOMAIN.  
FT DOMAIN 143 113 GLY-RICH.  
FT REPEAT 143 150 CXXCXGXG MOTIF.  
FT REPEAT 160 167 CXXCXGXG MOTIF.  
FT REPEAT 182 189 CXXCXGXG MOTIF.  
FT REPEAT 196 203 CXXCXGXG MOTIF.  
FT METAL 143 143 ZINC 1.

FT METAL 146 146 ZINC 1.  
FT METAL 160 160 ZINC 2.  
FT METAL 163 163 ZINC 2.  
FT METAL 182 182 ZINC 2.  
FT METAL 185 185 ZINC 2.  
FT METAL 196 196 ZINC 1.  
FT METAL 199 199 ZINC 1.  
FT HELIX 5 8  
FT TURN 9 10  
FT TURN 13 14  
FT HELIX 18 31  
FT TURN 32 35  
FT HELIX 36 38  
FT TURN 41 42  
FT HELIX 43 50  
FT TURN 51 51  
FT HELIX 52 56  
FT HELIX 59 64  
FT HELIX 65 67  
FT TURN 68 73  
FT TURN 131 133  
FT STRAND 140 142  
FT HELIX 144 146  
FT TURN 147 149  
FT STRAND 151 151  
FT STRAND 158 159  
FT TURN 161 165  
FT STRAND 168 173  
FT TURN 174 175  
FT STRAND 176 181  
FT TURN 183 187  
FT STRAND 190 191  
FT STRAND 195 195  
FT HELIX 197 199  
FT TURN 200 201  
FT STRAND 204 206  
SQ SEQUENCE 375 AA; 40969 MW; 8DE1863CC0BD8C3F CRC64;  
  
Query Match 100.0%; Score 86; DB 1; Length 375;  
Best Local Similarity 100.0%; Pred. No. 2e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 QKRAYDQYGHAAFEQ 16  
Db 60 QKRAYDQYGHAAFEQ 75  
  
RESULT 2  
DnaJ\_HAEDU STANDARD; PRT; 377 AA.  
ID DnaJ\_HAEDU  
AC P48208;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Chaperone protein dnaJ.  
GN DnaJ OR HD0188.  
OS Haemophilus ducreyi.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=730;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=35000HP / ATCC 700724;  
RA Parsons L.M.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=35000HP / ATCC 700724;  
RA Munson R.S. Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,  
RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;  
RT "The complete genome sequence of Haemophilus ducreyi.";  
Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: Acts as a co-chaperone. Stimulates, jointly with grpe,

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:29:38 ; Search time 20 Seconds  
(without alignments)  
76.953 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 OKRAAYDQYGHAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	376	1 HHECDJ	heat shock protein
2	86	100.0	376	2 G90630	DnaJ protein [lipo
3	86	100.0	376	2 G85481	chaperone with Dna
4	86	100.0	379	2 AB0058	chaperone protein
5	86	100.0	379	2 AP0503	DnaJ protein [lipo
6	86	100.0	394	2 C64112	heat shock protein
7	83	96.5	381	2 D82270	heat shock protein
8	82	95.3	392	2 S15295	molC protein - Rhi
9	75	87.2	377	2 TC5609	heat shock protein
10	75	87.2	377	2 F84947	DnaJ protein [lipo
11	73	84.9	375	2 B47042	heat shock protein
12	73	84.9	375	2 AC3502	chaperone protein
13	73	84.9	377	2 AD2591	molecular chaperon
14	73	84.9	377	2 D97373	chaperone protein
15	73	84.9	367	2 I40843	heat shock protein
16	64	74.4	370	2 C71729	DnaJ protein (dnaJ
17	64	74.4	376	2 E70361	chaperone DnaJ - A
18	63	73.3	190	2 B35388	heat shock protein
19	62	72.1	377	2 A83052	DnaJ protein PA476
20	62	72.1	377	2 H97728	DnaJ protein [lipo
21	60	69.8	379	2 A47079	heat shock protein
22	58	67.4	379	2 H86902	DnaJ protein [lipo
23	58	67.4	385	2 C87250	DnaJ protein [lipo
24	57	66.3	352	2 A49210	heat shock protein
25	57	66.3	364	2 D70164	heat shock protein
26	57	66.3	376	2 AD1621	heat shock protein
27	57	66.3	377	2 T43739	heat shock protein
28	57	66.3	377	2 AH1258	heat shock protein
29	57	66.3	377	2	

30	56	65.1	368	2 F82570	DnaJ protein XF233
31	56	65.1	372	2 H97928	heat-shock protein
32	56	65.1	378	2 D95060	DnaJ protein [lipo
33	53	61.6	499	2 G96831	hypothetical prote
34	52	60.5	372	2 B41874	heat shock protein
35	52	60.5	376	2 H69038	heat shock protein
36	52	60.5	389	2 S41748	heat shock protein
37	50	58.1	297	2 G81329	probable curved-DN
38	50	58.1	348	2 T04618	heat shock protein
39	50	58.1	379	2 B89939	DnaJ protein [lipo
40	50	58.1	383	2 B90603	heat shock protein
41	50	58.1	416	2 F71379	heat shock protein
42	49	57.0	335	2 T48161	heat shock protein
43	49	57.0	346	2 B84602	probable DnaJ prot
44	49	57.0	370	2 D83818	heat-shock protein
45	49	57.0	375	2 D82894	heat shock protein

## ALIGNMENTS

RESULT 1  
HHECDJ  
heat shock protein dnaJ - Escherichia coli (strain K-12)  
C/Species: Escherichia coli  
C/Date: 30-Jun-1988 #sequence\_revision 30-Jun-1988 #text\_change 01-Mar-2002  
C/Accession: A92572; A26298; S40537; G64721; A26299  
R/Bardwell, J.C.A.; Tilly, K.; Craig, E.; King, J.; Zylitz, M.; Georgopoulos, C.  
J. Biol. Chem. 261, 1782-1785, 1986  
A/Title: The nucleotide sequence of the Escherichia coli K12 dnaJ gene.  
A/Reference number: A92572; MUID:8611850; PMID:3003085  
A/Accession: A92572  
A/Molecule type: DNA  
A/Residues: 1-376 <BAR>  
A/Cross-references: GB:M12565; NID:G145767; PIDN:AAA23693.1; PID:G145769  
A/Experimental source: strain K12  
R/Ohki, M.; Tamura, F.; Nishimura, S.; Uchida, H.  
J. Biol. Chem. 261, 1778-1781, 1986  
A/Title: Nucleotide sequence of the Escherichia coli dnaJ gene and purification of the ge  
A/Reference number: A26298; MUID:8611849; PMID:3003084  
A/Accession: A26298  
A/Molecule type: DNA  
A/Residues: 1-376 <CHK>  
A/Cross-references: GB:D10483; GB:J01597; GB:J01683; GB:J01706; GB:K01298; GB:K01990; GB:  
R/Yura, T.; Mori, H.; Nagai, H.; Nagata, T.; Ishihama, A.; Fujita, N.; Isono, K.; Mizobu  
submitted to the EMBL Data Library, December 1992  
A/Description: Systematic sequencing of the Escherichia coli genome: analysis of the 0-2  
A/Reference number: S40537  
A/Accession: S40537  
A/Molecule type: DNA  
A/Residues: 1-376 <YUR>  
A/Cross-references: EMBL:D10483; NID:G216434; PIDN:BA01292.1; PID:G216441  
R/Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A/Title: The complete genome sequence of Escherichia coli K-12.  
A/Reference number: A64720; MUID:97426617; PMID:9278503  
A/Accession: G64721  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: DNA  
A/Residues: 1-376 <BLAT>  
A/Cross-references: GB:AE000112; GB:U00096; NID:G1786192; PIDN:AACT73126.1; PID:G1786197;  
A/Experimental source: strain K-12, substrain MG1655  
C/Comment: This protein is induced by heat shock under the control of the htrp gene prod  
C/Genetics:  
A/Gene: dnaJ  
A/Map position: 0 min  
C/Superfamily: heat shock protein dnaJ, dnaJ amino-terminal homology  
C/Keywords: DNA replication; heat shock; molecular chaperone; stress-induced protein  
F/5-70/Domain: dnaJ amino-terminal homology <DNJ>  
F/77-106/Region: G/F motif  
F/144-151/Region: CXXCXGXG repeat  
F/161-168/Region: CXXCXGXG repeat

F;183-190/Region: CXXCXGXG repeat  
F;197-204/Region: CXXCXGXG repeat

Query Match 100.0%; Score 86; DB 1; Length 376;  
Best Local Similarity 100.0%; Pred. No. 2.8e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
|||||  
Db 61 QKRAAYDQYGHAAFEQ 76

## RESULT 2

G90630

DnaJ protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)

C/Species: Escherichia coli

C/Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001

C/Accession: G90630

R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G. gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001

A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend

A/Reference number: A99629; MUID:21156231; PMID:11258796

A/Accession: G90630

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-376 <HAY>

A/Cross-references: GB:BA000007; PIDN:BA033438.1; PID:G13359471; GSPDB:GN00154

A/Experimental source: strain O157:H7, substrain RIMD 0509952

C/Genetics:

A/Gene: EGS0015

C/Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

## Query Match

Best Local Similarity 100.0%; Score 86; DB 2; Length 376;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
|||||  
Db 61 QKRAAYDQYGHAAFEQ 76

## RESULT 3

G85481

chaperone with DnaK, heat shock protein [imported] - Escherichia coli (strain O157:H7, S

C/Species: Escherichia coli

C/Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001

C/Accession: G85481

R/Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001

A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A/Reference number: A85480; MUID:21074935; PMID:11206551

A/Accession: G85481

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-376 <STO>

A/Cross-references: GB:AE005174; NID:G12512693; PIDN:AA054315.1; GSPDB:GN00145; UWGP:Z00

A/Experimental source: strain O157:H7, substrain EDL933

C/Genetics:

A/Gene: dnaJ

C/Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

## Query Match

Best Local Similarity 100.0%; Score 86; DB 2; Length 376;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
|||||  
Db 61 QKRAAYDQYGHAAFEQ 76

## RESULT 4

AB0058  
chaperone protein DnaJ [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis

C/Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001

C/Accession: AB0058

R/Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; F  
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, F  
Nature 413, 523-527, 2001

A/Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A/Reference number: AB0001; MUID:21470413; PMID:11586360

A/Accession: AB0058

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-379 <KUR>

A/Cross-references: GB:AL590842; PIDN:CAC89325.1; PID:G15978561; GSPDB:GN00175

C/Genetics:

A/Gene: dnaJ

C/Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

## Query Match

Best Local Similarity 100.0%; Score 86; DB 2; Length 379;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
|||||  
Db 61 QKRAAYDQYGHAAFEQ 76

## RESULT 5

AF0503

DnaJ protein [imported] - Salmonella enterica subsp. enterica serovar Typhi (strain CT18)

C/Species: Salmonella enterica subsp. enterica serovar Typhi

A/Note: this species has also been called Salmonella typhi

C/Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 18-Nov-2002

C/Accession: AF0503

R/Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
, S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001

A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;

A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov

A/Reference number: AB0502; MUID:21534947; PMID:11677608

A/Accession: AF0503

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-379 <PAR>

A/Cross-references: GB:AL513382; PIDN:CAD01166.1; PID:G16501296; GSPDB:GN00176

C/Genetics:

A/Gene: STY0013

C/Superfamily: heat shock protein dnaJ; dnaJ amino-terminal homology

## Query Match

Best Local Similarity 100.0%; Score 86; DB 2; Length 379;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16  
|||||  
Db 61 QKRAAYDQYGHAAFEQ 76

## RESULT 6

C64112

heat shock protein dnaJ - Haemophilus influenzae (strain Rd KW20)

C/Species: Haemophilus influenzae

C/Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 20-Aug-1999

C/Accession: C64112

R/Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, A.  
; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.  
, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.  
Science 269, 496-512, 1995

A/Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter, C.  
A/Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 10:05:13 ; Search time 54 Seconds  
(without alignments)  
83.718 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1:	geneseqp1980s:*
2:	geneseqp1990s:*
3:	geneseqp2000s:*
4:	geneseqp2001s:*
5:	geneseqp2002s:*
6:	geneseqp2003as:*
7:	geneseqp2003bs:*
8:	geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	86	100.0	16	5	AAU98852	Aau98852 E. Coli DN
2	86	100.0	16	6	ABR55126	AbR55126 E. coli d
3	86	100.0	131	3	AAB11396	Aab11396 E. coli e
4	86	100.0	131	4	AAB74197	Aab74197 OmpA-J do
5	86	100.0	131	4	AAB70767	Aab70767 Expressio
6	86	100.0	131	4	AAV72018	AAV72018 E. coli O
7	86	100.0	340	4	AAU17588	Aau17588 Novel sig
8	86	100.0	340	7	ADB94296	AdB94296 Human nov
9	86	100.0	341	4	AAU17587	Aau17587 Novel sig
10	86	100.0	341	7	ADB94295	AdB94295 Human nov
11	86	100.0	373	6	ABM67485	Abm67485 Photornab
12	86	100.0	399	3	AAB11395	Aab11395 E. coli e
13	86	100.0	399	4	AAB74196	Aab74196 OmpA-DnaJ
14	86	100.0	399	4	AAB70766	Expressio
15	86	100.0	399	4	AAV72017	E. coli O
16	86	100.0	476	4	ABG17771	Abg17771 Novel hum
17	81	94.2	15	2	AAW25795	Aaw25795 Peptide d
18	81	94.2	15	3	AAV79543	Aay79543 E. coli d
19	81	94.2	15	4	AAW99341	Aam99341 Vaccine r
20	81	94.2	15	5	AAU09842	Aau09842 E. coli dn
21	81	94.2	15	5	AAU98876	Aau98876 E. coli D
22	81	94.2	15	5	AAE19457	Aae19457 Human hea
23	81	94.2	15	6	ABR55132	AbR55132 E. coli D
24	81	94.2	16	4	AAW99344	Aam99344 Vaccine r
25	81	94.2	118	4	AAU02075	Aau02075 Synthetic

26	81	94.2	459	4	AAU02077	Aau02077 Synthetic
27	73	84.9	344	6	ABP80877	Abp80877 N. gonorr
28	72	83.7	15	2	AAR95446	Aar95446 RA suscep
29	72	83.7	15	2	AAW25796	Aaw25796 Peptide d
30	72	83.7	15	5	AAE19458	Aae19458 Heat shoc
31	72	83.7	15	6	ABR55133	AbR55133 E. coli d
32	64.5	75.0	14	2	AAR95445	Aar95445 RA suscep
33	63	73.3	332	3	AAG14804	Aag14804 Arabidops
34	63	73.3	367	3	AAG14803	Aag14803 Arabidops
35	63	73.3	456	3	AAG14802	Aag14802 Arabidops
36	61	70.9	385	6	ADB10662	AdB10662 Alloiococ
37	60	69.8	368	3	AAW05934	Aaw05934 Protein d
38	59	68.6	332	3	AAG48466	Aag48466 Arabidops
39	59	68.6	367	3	AAG48465	Aag48465 Arabidops
40	59	68.6	456	3	AAG48464	Aag48464 Arabidops
41	58	67.4	379	5	ABB55579	Abb55579 Lactococc
42	57	66.3	377	5	ABB48799	Abb48799 Listeria
43	56	65.1	352	2	AAW22358	Aaw22358 S. pneumo
44	56	65.1	378	6	ABU00882	Abu00882 S. pneumo
45	56	65.1	378	6	ABP81556	Abp81556 Streptoco

ALIGNMENTS

RESULT 1	
AAU98852	AAU98852 standard; peptide; 16 AA.
AC	AAU98852;
XX	
XX	
AC	
XX	
DT	22-AUG-2002 (first entry)
XX	
DE	E.Coli DNAJ 61 immunogenic peptide.
XX	
KW	Immunogenic peptide; heat shock protein; HSP; DNAJ; immunomodulatory;
KW	cytostatic; antiinflammatory; antibacterial; antiarthritic;
KW	autoimmune disease; arthritis; articular juvenile idiopathic arthritis;
KW	infectious disease; inflammatory bowel disease; cancer;
KW	mucosal tolerisation; DNA vaccination; anergy induction.
XX	
OS	Escherichia coli.
XX	
PN	WO200236611-A2.
XX	
PD	10-MAY-2002.
XX	
PF	31-OCT-2001; 2001WO-US045344.
XX	
PR	01-NOV-2000; 2000US-0245181P.
XX	
PA	(REGC ) UNIV CALIFORNIA.
PA	(MART/) MARTINI A.
XX	
PI	Martini A, Albani S, Carson DA, Prakken BJ;
XX	
DR	WPI; 2002-489999/52.
XX	
PT	New immunomodulatory peptides from heat shock proteins, useful for
PT	treating immunological disorder in subjects such as humans, e.g.
PT	autoimmune disease (e.g. arthritis), infectious disease, inflammatory
PT	bowel disease or cancer.
XX	
PS	Claim 4; Page 55; 84pp; English.
XX	
CC	This invention relates to the use of a peptide, which is an immunogenic
CC	portion derived from a dnaJ heat shock protein (hsp) in modulating an
CC	immune response in a subject. The peptides of the invention may have
CC	immunomodulatory, cytostatic, antiinflammatory, antibacterial or
CC	antiarthritic properties, and can stimulate expression of interleukins,
CC	tumour necrosis factor and transforming growth factor beta. The
CC	immunogenic peptide is useful for modulating (i.e. augmenting/inducing or
CC	reducing/inhibiting) an immune response in a subject having an

CC immunological disorder (e.g. autoimmune disease such as arthritis or  
CC articular juvenile idiopathic arthritis), an infectious disease, an  
CC inflammatory bowel disease or cancer. The immunogenic peptide of the  
CC invention is also useful for modulating immunoeffector cell  
CC responsiveness in a subject. The immunogenic peptide is particularly  
CC useful for treating the above-mentioned diseases in mammals, e.g. cat,  
CC dog, horse, farm animal (e.g. ovine, bovine or porcine) or human. In  
CC general, the peptide is useful in methods involving mucosal tolerisation,  
CC DNA vaccination, anergy induction or active immunisation. The present  
CC sequence represents an E. coli DNAJ immunogenic peptide of the invention  
XX  
SQ Sequence 16 AA;

Query Match 100.0%; Score 86; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 4.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QKRAAYDQYGHAAFEQ 16  
|||  
Db 1 QKRAAYDQYGHAAFEQ 16

RESULT 2  
ABR55126  
ID ABR55126 standard; peptide; 16 AA.

AC ABR55126;

DT 03-JUL-2003 (first entry)

DE E. coli dnaJ61 antigen-specific epitope peptide.

XX Antigen-specific epitope; immune response; T cell; cytokine;  
KM antiarthritic; antidiabetic; neuroprotective; anti-inflammatory;  
KW cytostatic; antithyroid; anti-asthmatic; immunosuppressive;  
KW antipsoriatic; anti-ulcer; antianaemic; cardiant; respiratory;  
KW antiallergic; dermatological; antipsoriatic.

OS Escherichia coli.

PN W02003026579-A2.

PD 03-APR-2003.

PF 25-SEP-2002; 2002WO-US030578.

PR 25-SEP-2001; 2001US-0325499P.

PR 11-DEC-2001; 2001US-0339284P.

PA (REGC ) UNIV CALIFORNIA.

PI Albani S, Martins A;

WPI; 2003-430097/40.

XX Modulating an immune response in a subject having an immune-related  
PT disorder, e.g. arthritis by administering an antigen-specific epitope and  
PT a cytokine or an agent that effects cytokine activity or expression.  
XX  
PS Disclosure; Page 9; 41pp; English.

XX The invention relates to a novel method for modulating an immune response  
CC in a subject having an immune-related disorder. The method comprises: (a)  
CC administering an antigen-specific epitope; where administration provides  
CC epitope-specific T cell immune modulation; and (b) administering a  
CC cytokine, an agent that effects cytokine activity or expression, or an  
CC anticytokine therapy. The method of the invention has antiarthritic,  
CC antidiabetic, neuroprotective, anti-inflammatory, cytostatic,  
CC antithyroid, anti-asthmatic, immunosuppressive, antipsoriatic, anti-  
CC ulcer, antianaemic, cardiant, respiratory general, antiallergic, anti-  
CC dermatological, and antipsoriatic activity. The method is useful for  
CC modulating an immune response in a subject having an immune-related  
CC disorder. The present sequence is used in the exemplification of the

CC invention  
XX  
SQ Sequence 16 AA;

Query Match 100.0%; Score 86; DB 6; Length 16;  
Best Local Similarity 100.0%; Pred. No. 4.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QKRAAYDQYGHAAFEQ 16  
|||  
Db 1 QKRAAYDQYGHAAFEQ 16

RESULT 3  
AAB11396  
ID AAB11396 standard; protein; 131 AA.

AC AAB11396;

DT 22-FEB-2001 (first entry)

DE E. coli expression plasmid PUBS520-PIN-J-Domain encoded protein.

XX Eukaryotic protein; protease; interferon; antibody; hormone;  
KW disulfide bridge.

OS Escherichia coli.

OS Synthetic.

PN EP1048732-A1.

PD 02-NOV-2000.

PF 26-APR-1999; 99EP-00107412.

PR 26-APR-1999; 99EP-00107412.

PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

WPI; 2000-674185/66.

DR N-PSDB; AAC66072.

XX Preparation of water-soluble eukaryotic polypeptides with disulfide  
PT bridges e.g. rPA, comprises cultivation of prokaryotic cells in the  
PT presence of arginine or amide compound.  
XX  
PS Example 3; Page 17; 40pp; German.

XX This invention describes a novel preparation of a water-soluble,  
CC naturally occurring eukaryotic polypeptide containing two or more  
CC cysteine units bound via a disulfide bridge which comprises cultivation  
CC of prokaryotic cells in the presence of arginine or an amide compound.  
CC The method is useful for the preparation of eukaryotic proteins e.g.  
CC proteases, interferons, protein hormones, antibodies or antibody  
CC fragments (e.g. a single chain FV fragment that binds to thyroid  
CC stimulating hormone). It is especially useful for preparing proteins with  
CC more than five disulfide bridges, e.g. recombinant plasminogen activator  
CC (rPA). The technique is simple and does not require in vitro after-  
CC treatment, such as the removal of inclusion bodies, reduction or  
CC naturalization  
XX  
SQ Sequence 131 AA;

Query Match 100.0%; Score 86; DB 3; Length 131;  
Best Local Similarity 100.0%; Pred. No. 4e-07;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 QKRAAYDQYGHAAFEQ 16  
|||  
Db 84 QKRAAYDQYGHAAFEQ 99

RESULT 4

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:41:15 ; Search time 44 Seconds

(without alignments)  
113.425 Million cell updates/sec

Title: US-10-001-938-3

Perfect score: 86

Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62

Searched: 1279676 seqs, 311918243 residues

Total number of hits satisfying chosen parameters: 218390

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	86	100.0	16	US-10-001-938-3	Sequence 3, Appl1
2	81	94.2	15	US-09-756-983-10	Sequence 10, Appl1
3	81	94.2	15	US-09-828-574-10	Sequence 10, Appl1
4	81	94.2	15	US-10-001-938-27	Sequence 27, Appl1
5	81	94.2	15	US-10-299-540-4	Sequence 4, Appl1
6	81	94.2	15	US-10-299-184-4	Sequence 4, Appl1
7	81	94.2	15	US-10-239-313A-444	Sequence 444, App
8	81	94.2	16	US-10-239-313A-447	Sequence 447, App
9	72	83.7	15	US-09-828-574-11	Sequence 11, Appl1
10	72	83.7	15	US-10-299-540-5	Sequence 5, Appl1
11	72	83.7	15	US-10-299-184-5	Sequence 5, Appl1
12	54	62.8	15	US-10-299-540-6	Sequence 6, Appl1
13	54	62.8	15	US-10-299-184-6	Sequence 6, Appl1
14	34	39.5	9	US-10-283-423-154	Sequence 154, App
15	34	39.5	9	US-10-283-423-155	Sequence 155, App

16	34	39.5	9	14	US-10-283-423-157	Sequence 157, App
17	34	39.5	9	14	US-10-283-423-158	Sequence 158, App
18	34	39.5	9	14	US-10-283-423-159	Sequence 159, App
19	34	39.5	9	14	US-10-213-821-154	Sequence 154, App
20	34	39.5	9	14	US-10-213-821-155	Sequence 155, App
21	34	39.5	9	14	US-10-213-821-157	Sequence 157, App
22	34	39.5	9	14	US-10-213-821-158	Sequence 158, App
23	34	39.5	9	14	US-10-213-821-159	Sequence 159, App
24	34	39.5	9	16	US-10-736-048-154	Sequence 154, App
25	34	39.5	9	16	US-10-736-048-155	Sequence 155, App
26	34	39.5	9	16	US-10-736-048-157	Sequence 157, App
27	34	39.5	9	16	US-10-736-048-158	Sequence 158, App
28	34	39.5	9	16	US-10-736-048-159	Sequence 159, App
29	34	39.5	14	14	US-10-161-959-6	Sequence 6, Appl1
30	34	39.5	14	14	US-10-283-423-160	Sequence 160, App
31	34	39.5	14	14	US-10-213-821-160	Sequence 160, App
32	34	39.5	14	16	US-10-736-048-160	Sequence 160, App
33	34	39.5	15	9	US-09-756-983-11	Sequence 11, Appl1
34	34	39.5	15	14	US-10-161-959-5	Sequence 5, Appl1
35	34	39.5	15	14	US-10-299-540-7	Sequence 7, Appl1
36	34	39.5	15	14	US-10-299-540-8	Sequence 8, Appl1
37	34	39.5	15	14	US-10-299-184-7	Sequence 7, Appl1
38	34	39.5	15	14	US-10-299-184-8	Sequence 8, Appl1
39	34	39.5	15	14	US-10-239-313A-443	Sequence 443, App
40	32	37.2	13	15	US-10-089-887-58	Sequence 58, Appl1
41	31	36.0	14	14	US-10-161-959-28	Sequence 28, Appl1
42	31	36.0	15	12	US-09-841-091B-10	Sequence 10, Appl1
43	31	36.0	15	12	US-10-056-583-30	Sequence 30, Appl1
44	31	36.0	15	14	US-10-161-959-10	Sequence 10, Appl1
45	31	36.0	15	14	US-10-251-703-10	Sequence 10, Appl1

ALIGNMENTS

RESULT 1

US-10-001-938-3

; Sequence 3, Application US/10001938

; Publication No. US20030031679A1

; GENERAL INFORMATION:

; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA

; APPLICANT: ALBANI, Salvatore

; APPLICANT: CARSON, Dennis

; APPLICANT: PRAKEN, Berent

; APPLICANT: MARTINI, Alberto

; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES DERIVED FROM HEAT SHOCK PROTEINS AND US1

; FILE REFERENCE: UCSD1360-1

; CURRENT APPLICATION NUMBER: US/10/001,938

; CURRENT FILING DATE: 2001-10-31

; PRIOR APPLICATION NUMBER: US 60/245,181

; PRIOR FILING DATE: 2000-11-01

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 3

; LENGTH: 16

; TYPE: PRT

; ORGANISM: Escherichia coli

US-10-001-938-3

Query Match 100.0%; Score 86; DB 14; Length 16;

Best Local Similarity 100.0%; Pred. No. 9.3e-08;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFEQ 16

Db 1 QKRAAYDQYGHAAFEQ 16

RESULT 2

US-09-756-983-10

; Sequence 10, Application US/09756983

; Patent No. US20020122818A1

GENERAL INFORMATION:  
; APPLICANT: Albani, Salvatore  
; TITLE OF INVENTION: METHODS FOR ISOLATION, QUANTIFICATION,  
; TITLE OF INVENTION: CHARACTERIZATION AND MODULATION OF  
; TITLE OF INVENTION: ANTIGEN-SPECIFIC T CELLS  
; FILE REFERENCE: 246/285-CIP  
; CURRENT APPLICATION NUMBER: US/09/756,983  
; CURRENT FILING DATE: 2001-01-09  
; PRIOR APPLICATION NUMBER: 60/105,018  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 09/421,506  
; PRIOR FILING DATE: 1999-10-19  
; PRIOR APPLICATION NUMBER: PCT/US99/2466  
; PRIOR FILING DATE: 1999-10-19  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: E. coli  
; FEATURE:  
; OTHER INFORMATION: dnaJp1 heat shock protein  
US-09-756-983-10

Query Match 94.2%; Score 81; DB 9; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAFFE 15  
Db 1 OKRAAYDQYGHAFFE 15

RESULT 3  
US-09-828-574-10  
; Sequence 10, Application US/09828574  
; Patent No. US20020146759A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: ALBANI, Salvatore  
; APPLICANT: PRAKEN, Berent J.  
; TITLE OF INVENTION: STRESS PROTEINS AND PEPTIDES AND METHODS OF USE THEREOF  
; FILE REFERENCE: UCSD1310-1  
; CURRENT APPLICATION NUMBER: US/09/828,574  
; CURRENT FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: US 60/224,104  
; PRIOR FILING DATE: 2000-08-09  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: dnaJp1 peptide  
US-09-828-574-10

Query Match 94.2%; Score 81; DB 9; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAFFE 15  
Db 1 OKRAAYDQYGHAFFE 15

RESULT 4  
US-10-001-938-27  
; Sequence 27, Application US/10001938  
; Publication No. US20030031679A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: ALBANI, Salvatore

APPLICANT: CARSON, Dennis  
; APPLICANT: PRAKEN, Berent  
; APPLICANT: MARTINI, Alberto  
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES DERIVED FROM HEAT SHOCK PROTEINS AND USE  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: UCSD1360-1  
; CURRENT APPLICATION NUMBER: US/10/001,938  
; CURRENT FILING DATE: 2001-10-31  
; PRIOR APPLICATION NUMBER: US 60/245,181  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 27  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Escherichia coli  
US-10-001-938-27

Query Match 94.2%; Score 81; DB 14; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAFFE 15  
Db 1 OKRAAYDQYGHAFFE 15

RESULT 5  
US-10-299-540-4  
; Sequence 4, Application US/10299540  
; Publication No. US20030143238A1  
; GENERAL INFORMATION:  
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA  
; APPLICANT: CARSON, Dennis A.  
; APPLICANT: ALBANI, Salvatore  
; TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN INDUCING IMMUNE PROTECT  
; TITLE OF INVENTION: AGAINST ARTHRITIC PEPTIDES INVOLVED IN THE PATHOGENESIS OF R  
; TITLE OF INVENTION: ARTHRITIS  
; FILE REFERENCE: UCSD1370-7  
; CURRENT APPLICATION NUMBER: US/10/299,540  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: US 09/616,247  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/107,615  
; PRIOR FILING DATE: 1998-06-30  
; PRIOR APPLICATION NUMBER: US 08/618,464  
; PRIOR FILING DATE: 1996-03-15  
; PRIOR APPLICATION NUMBER: US 08/246,988  
; PRIOR FILING DATE: 1994-05-20  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic dnaJp1  
US-10-299-540-4

Query Match 94.2%; Score 81; DB 14; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.1e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 OKRAAYDQYGHAFFE 15  
Db 1 OKRAAYDQYGHAFFE 15

RESULT 6  
US-10-299-184-4  
; Sequence 4, Application US/10299184  
; Publication No. US20030147910A1  
; GENERAL INFORMATION:



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:38:55 ; Search time 20 Seconds  
(without alignments)  
76.953 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 2772

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	34	39.5	11	2	A60656	perisulfakinin - A
2	34	39.5	14	2	A56632	neosulfakinin-II -
3	31	36.0	10	1	GMROL2	leucosulfakinin-II
4	31	36.0	10	2	B60656	leucosulfakinin II
5	31	36.0	13	2	S47372	T-cell antigen rec
6	31	36.0	13	2	S47384	T-cell antigen rec
7	30	34.9	11	1	GMROL	leucosulfakinin -
8	30	34.9	13	2	S47368	T-cell antigen rec
9	28	32.6	13	2	S47359	T-cell antigen rec
10	28	32.6	13	2	S47365	T-cell antigen rec
11	28	32.6	13	2	S47374	T-cell antigen rec
12	26	30.2	11	2	PT0214	T-cell receptor be
13	26	30.2	13	2	S47388	T-cell antigen rec
14	26	30.2	13	2	PH1479	T-cell receptor be
15	26	30.2	14	4	I52618	hemoglobin beta ch
16	26	30.2	15	2	PH0770	T-cell receptor be
17	25	29.1	13	2	S32471	lymadFamide 1 - g
18	25	29.1	14	2	PH1586	Ig H chain V-D-J r
19	24	27.9	16	2	PH1622	Ig H chain V-D-J r
20	24	27.9	16	2	PH1589	Ig H chain V-D-J r
21	23	26.7	8	2	S11078	glucose-6-phosphat
22	23	26.7	13	2	S32472	lymadFamide 2 - g
23	23	26.7	13	2	A60379	factor X activator
24	22	25.6	13	2	S32473	lymadFamide 3 - g
25	22	25.6	13	2	S47383	T-cell antigen rec
26	22	25.6	15	2	PH1319	Ig heavy chain DJ
27	22	25.6	15	2	F28587	T-cell receptor be
28	22	25.6	15	2	I53284	T-cell receptor be
29	21	24.4	10	2	PT0215	T-cell receptor be

30	21	24.4	12	2	PH0746	T-cell receptor be
31	21	24.4	12	2	PH1461	T-cell receptor be
32	21	24.4	12	2	PH0771	T-cell receptor be
33	21	24.4	13	2	S32474	lymadFamide 4 - g
34	21	24.4	13	2	S47361	T-cell antigen rec
35	21	24.4	13	2	I51905	collecting duct wa
36	21	24.4	14	2	S48685	extension protein
37	21	24.4	14	2	G44957	photosystem II oxy
38	21	24.4	14	2	PH0755	T-cell receptor be
39	21	24.4	14	2	A17150	glucose 1-dehydrog
40	21	24.4	15	2	PC4269	unidentified QR310
41	20	23.3	10	2	PT0309	Ig heavy chain CRD
42	20	23.3	13	2	G22565	R-phycocerythrin ga
43	20	23.3	13	2	PN0168	phosphopyruvate hy
44	20	23.3	14	2	PA0013	photosystem II oxy
45	20	23.3	14	2	PT0232	Ig heavy chain CRD

ALIGNMENTS

RESULT 1  
A60656  
perisulfakinin - American cockroach  
C:Species: Periplaneta americana (American cockroach)  
C>Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 11-Jul-1997  
C:Accession: A60656  
R:Veenstra, J.A.  
Neuropeptides 14, 145-149, 1989  
A:Title: Isolation and structure of two gastrin/CCK-like neuropeptides from the American  
A:Reference number: A60656; MUID:90137190; PMID:2615921  
A:Accession: A60656  
A:Molecule type: protein  
A:Residues: 1-11 <VEE>  
C:Comment: This neuropeptide stimulates hindgut contractions.  
C:Keywords: amidated carboxyl end; neuropeptide; sulfoprotein  
F:6/Binding site: sulfate (Tyr) (covalent) #status experimental  
F:11/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 39.5%; Score 34; DB 2; Length 11;  
Best Local Similarity 55.6%; Pred. No. 9;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 6 YDQYGHAAF 14  
Db 3 FDDYGHMRF 11

RESULT 2  
A56632  
neosulfakinin-II - flesh fly (Sarcophaga bullata)  
N:Alternate names: Neb-SK-II  
N:Contains: neosulfakinin-I (Neb-SK-I)  
C:Species: Sarcophaga bullata  
C>Date: 21-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 20-Jun-2000  
C:Accession: A56632  
R:Fonagy, A.; Schoofs, L.; Proost, P.; Van Damme, J.; De Loof, A.  
Comp. Biochem. Physiol. C 103, 135-142, 1992  
A:Title: Isolation and primary structure of two sulfakinin-like peptides from the fleshf  
A:Reference number: A56632; MUID:93083101; PMID:1360367  
A:Accession: A56632  
A:Molecule type: protein  
A:Residues: 1-14 <FON>  
A:Experimental source: heads  
A:Note: sequence extracted from NCBI backbone (NCBIP:120391)  
C:Keywords: amidated carboxyl end; neuropeptide; sulfoprotein  
F:1-14/Product: neosulfakinin-II #status experimental <NSK2>  
F:6-14/Product: neosulfakinin-I #status experimental <NSK1>  
F:9/Binding site: sulfate (Tyr) (covalent) #status predicted  
F:14/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 39.5%; Score 34; DB 2; Length 14;  
Best Local Similarity 55.6%; Pred. No. 12;

Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 YDYGHAAF 14  
: ||| |  
|  
Db 6 FDDYGHMRF 14

## RESULT 3

GMROL2  
leucosulfakinin-II - Madeira cockroach  
N:Alternate names: LSK-II  
C:Species: Leucophaea maderae (Madeira cockroach)  
C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 13-Sep-1996  
C:Accession: A26335  
R:Nachman, R.J.; Holman, G.M.; Cook, B.J.; Haddon, W.F.; Ling, N.  
Biochem. Biophys. Res. Commun. 140, 357-364, 1986  
A:Title: Leucosulfakinin-II, a blocked sulfated insect neuropeptide with homology to chc  
A:Reference number: A26335; MUID:87048769; PMID:3778455  
A:Accession: A26335  
A:Molecule type: protein  
A:Residues: 1-10 <NAC>  
C:Comment: This peptide was isolated from head extracts. It stimulates muscle contractic  
C:Superfamily: gastrin  
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid; sulfoprotein  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:5/Binding site: sulfate (Tyr) (covalent) #status experimental  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 36.0%; Score 31; DB 1; Length 10;  
Best Local Similarity 62.5%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 DQYGHAAF 14  
: ||| |  
|  
Db 3 DDYGHMRF 10

## RESULT 4

B60656  
leucosulfakinin II, non-sulfated - American cockroach  
C:Species: Periplaneta americana (American cockroach)  
C:Date: 14-May-1993 #sequence\_revision 14-May-1993 #text\_change 08-Dec-1995  
C:Accession: B60656  
R:Veestra, J.A.  
Neuropeptides 14, 145-149, 1989  
A:Title: Isolation and structure of two gastrin/CK-like neuropeptides from the American  
A:Reference number: A60656; MUID:90137190; PMID:2615921  
A:Accession: B60656  
A:Molecule type: protein  
A:Residues: 1-10 <VE>  
C:Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:10/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 36.0%; Score 31; DB 2; Length 10;  
Best Local Similarity 62.5%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 7 DQYGHAAF 14  
: ||| |  
|  
Db 3 DDYGHMRF 10

## RESULT 5

S47372  
T-cell antigen receptor VJ junction beta chain - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S47372  
R:Lehner, P.J.  
submitted to the EMBL Data Library, August 1994  
A:Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T c  
ference number: S47355

A:Accession: S47372  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-13 <LEH>  
A:Cross-references: EMBL:Z35697; NID:g527485; PIDN:CAA84766.1; PID:g527486  
C:Keywords: T-cell receptor

Query Match 36.0%; Score 31; DB 2; Length 13;  
Best Local Similarity 71.4%; Pred. No. 36;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RAYDQY 9  
: ||| |  
|  
Db 6 RSAYEQY 12

## RESULT 6

S47384  
T-cell antigen receptor VJ junction beta chain - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 05-Nov-1999  
C:Accession: S47384  
R:Lehner, P.J.  
submitted to the EMBL Data Library, August 1994  
A:Description: Human HLA-A0201 restricted recognition of influenza A is dominated by T c  
A:Reference number: S47355  
A:Accession: S47384  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-13 <LEH>  
A:Cross-references: EMBL:Z35704; NID:g527501; PIDN:CAA84773.1; PID:g527502  
C:Keywords: T-cell receptor

Query Match 36.0%; Score 31; DB 2; Length 13;  
Best Local Similarity 71.4%; Pred. No. 36;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 RAYDQY 9  
: ||| |  
|  
Db 6 RSAYEQY 12

## RESULT 7

GMROL  
leucosulfakinin - Madeira cockroach  
N:Alternate names: LSK  
C:Species: Leucophaea maderae (Madeira cockroach)  
C:Date: 17-Mar-1987 #sequence\_revision 17-Mar-1987 #text\_change 13-Sep-1996  
C:Accession: A01622  
R:Nachman, R.J.; Holman, G.M.; Haddon, W.F.; Ling, N.  
Science 234, 71-73, 1986  
A:Title: leucosulfakinin, a sulfated insect neuropeptide with homology to gastrin and chc  
A:Reference number: A01622; MUID:86315858; PMID:3749893  
A:Accession: A01622  
A:Molecule type: protein  
A:Residues: 1-11 <NAC>  
C:Superfamily: gastrin  
C:Keywords: amidated carboxyl end; hormone; sulfoprotein  
F:6/Binding site: sulfate (Tyr) (covalent) #status experimental  
F:11/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 34.9%; Score 30; DB 1; Length 11;  
Best Local Similarity 44.4%; Pred. No. 45;  
Matches 4; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 YDYGHAAF 14  
: ||| |  
|  
Db 3 FDDYGHMRF 11

## RESULT 8

S47368  
T-cell antigen receptor VJ junction beta chain - human

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:34:09 ; Search time 11 Seconds  
(without alignments)

75.738 Million cell updates/sec

Title: US-10-001-938-3

Perfect score: 86

Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 880

Minimum DB seq length: 0

Maximum DB seq length: 16

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34	39.5	9	1	NSK1_SARBU
2	34	39.5	11	1	LSKP_PERAM
3	34	39.5	14	1	NSK2_SARBU
4	33	38.4	12	1	LOSK_LOCOMI
5	31	36.0	10	1	LSK2_LEUMA
6	30	34.9	11	1	LSK1_LEUMA
7	25	29.1	13	1	NP1_LYMST
8	23	26.7	13	1	NP2_LYMST
9	22	25.6	13	1	NP3_LYMST
10	21	24.4	13	1	NP4_LYMST
11	21	24.4	13	1	PSBP_PINPS
12	21	24.4	15	1	ECDA_LYMDI
13	19	22.1	12	1	NUDM_CANFA
14	19	22.1	13	1	NP5_LYMST
15	19	22.1	14	1	MY14_MAIZE
16	19	22.1	15	1	UC14_MAIZE
17	19	22.1	16	1	ARCD_PSEPU
18	18	20.9	7	1	UN06_PINPS
19	18	20.9	12	1	PSP3_PHYPA
20	18	20.9	12	1	RS19_CLYEP
21	18	20.9	12	1	RS19_ELYEP
22	18	20.9	14	1	ADFA_TENMO
23	18	20.9	14	1	RS19_CLOPP
24	18	20.9	14	1	RS19_LOWBP
25	18	20.9	15	1	ACT_FINPS
26	18	20.9	15	1	FIBA_ANAPL
27	18	20.9	15	1	UC19_MAIZE
28	18	20.9	16	1	ALRX_PSEPU
29	18	20.9	16	1	MLB_SQUAC
30	18	20.9	16	1	RIPK_TRIKI
31	18	20.9	16	1	SAL3_ONCMY
32	17	19.8	14	1	TKN1_SCHGR
33	17	19.8	15	1	RS10_BACST

34	17	19.8	15	1	SAL1_ONCMY	P81369 oncorhynchu
35	17	19.8	15	1	SODV_ENTAE	P22799 enterobacte
36	17	19.8	15	1	UC23_MAIZE	P80629 zea mays (m
37	17	19.8	16	1	ALL1_MAIZE	P41839 calliphora
38	17	19.8	16	1	PPAC_BACME	P56948 bacillus me
39	16	18.6	8	1	LCK1_LEUMA	P21140 leucophaea
40	16	18.6	8	1	NS3_MYCTU	P81152 mycobacteri
41	16	18.6	10	1	AEGL_AGRAE	P83465 agrocye ae
42	16	18.6	10	1	FAR2_PENMO	P83317 penaeus mon
43	16	18.6	10	1	RL16_ACHLA	P29221 acholeplasm
44	16	18.6	10	1	TKN1_SCYCA	P08608 scylliorhinu
45	16	18.6	11	1	CA32_LITCI	P82090 litorea cit

ALIGNMENTS

RESULT 1	NSK1_SARBU	STANDARD;	PRT;	9 AA.
ID	NSK1_SARBU			
AC	P41492;			
DT	01-NOV-1995 (Rel. 32, Created)			
DT	01-NOV-1995 (Rel. 32, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	Neosulfakinin-I (NEB-SK-I).			
OS	Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;			
OC	Sarcophagidae; Sarcophaga.			
OX	NCBI_TaxID=7385;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Head;			
RX	MEDLINE=93083101; PubMed=1360367;			
RA	Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;			
RT	"Isolation and primary structure of two sulfakinin-like peptides from			
RT	the fleshfly, Neobellieria bullata.";			
RL	Comp. Biochem. Physiol. 103C:135-142(1992).			
CC	-1- FUNCTION: Myotropic peptide.			
CC	-1- SIMILARITY: Belongs to the gastrin/cholecystokinin family.			
DR	InterPro; IPR001651; Gastrin.			
DR	PROSITE; PS00259; GASTRIN; 1.			
KW	Neuropeptide; Amidation; Sulfation.			
FT	MOD_RES 4 4 SULFATION (POTENTIAL).			
FT	MOD_RES 9 9 AMIDATION (POTENTIAL).			
FT	SEQUENCE 9 AA; 1187 MW; 8B0A0691E86B5AAA CRC64;			
QY	6 YDQYGHAAF 14			
Db	1 FDDYGHMRF 9			
RESULT 2	LSKP_PERAM	STANDARD;	PRT;	11 AA.
ID	LSKP_PERAM			
AC	P36885;			
DT	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	Periplaneta americana (Pea-SK-I).			
OS	Periplaneta americana (American cockroach).			
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;			
OC	Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;			
OC	Blattidae; Periplaneta.			
OX	NCBI_TaxID=6978;			
RN	[1]			
RP	SEQUENCE.			
RC	TISSUE=Corpora cardiaca;			
RX	MEDLINE=90137190; PubMed=2615921;			

```

RA Veenstra J.A.;
RT "Isolation and structure of two gastrin/CCK-like neuropeptides from
RL the American cockroach homologous to the leucosulfakinins.";
CC -!- FUNCTION: Stimulates hindgut contractions.
CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR PIR; A60656; A60656.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Hormone; Amidation; Sulfation.
FT MOD_RES 6 11
FT MOD_RES 11 11
SQ SEQUENCE 11 AA; 1445 MW; 8B4E0680E86B5AAA CRC64;

Query Match 39.5%; Score 34; DB 1; Length 11;
Best Local Similarity 55.6%; Pred. No. 5;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 YDQYGHAAF 14
Db 3 FDDYGHMRF 11

RESULT 3
NSK2_SARBU STANDARD; PRT; 14 AA.
ID _NSK2_SARBU
AC P41493;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Neosulfakinin-II (NEB-SK-II).
OS Sarcophaga bullata (Grey flesh fly) (Neobellieria bullata).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC Sarcophagidae; Sarcophaga.
CX NCBI_TaxID=7385;
RN [1]
RP SEQUENCE.
RC TISSUE=Head;
RX MEDLINE=93083101; PubMed=1360367;
RA Fonagy A., Schoofs L., Proost P., van Damme J., de Loof A.;
RT "Isolation and primary structure of two sulfakinin-like peptides from
RL the fleshfly, Neobellieria bullata.";
RL Comp. Biochem. Physiol. 103C:135-142(1992).
CC -!- FUNCTION: Myotropic peptide.
CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR PIR; A56632; A56632.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Neuropeptide; Amidation; Sulfation.
FT MOD_RES 9 9
FT MOD_RES 14 14
FT MOD_RES 14 14
SQ SEQUENCE 14 AA; 1794 MW; 8B4E06D5B61C62AA CRC64;

Query Match 39.5%; Score 34; DB 1; Length 14;
Best Local Similarity 55.6%; Pred. No. 6.5;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 6 YDQYGHAAF 14
Db 6 FDDYGHMRF 14

RESULT 4
LOSK_LOCMI STANDARD; PRT; 12 AA.
ID _LOSK_LOCMI
AC P47733;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Sulfakinin (LOM-SK).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

```

```

OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Brain;
RA Schoofs L., Holman G.L., Hayes T.K., Nachman R.J., de Loof A.;
RL (In) McCaffery A., Wilson I. (eds.);
RL Chromatography and Isolation of insect hormones and pheromones,
RL pp.231-241, Plenum Press, New York (1990).
CC -!- FUNCTION: Myotropic peptide.
CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Hormone; Amidation; Sulfation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1
FT MOD_RES 7 7
FT MOD_RES 7 7
FT MOD_RES 12 12
FT MOD_RES 12 12
SQ SEQUENCE 12 AA; 1440 MW; 9B5B5DA9BD6B5AAA CRC64;

Query Match 38.4%; Score 33; DB 1; Length 12;
Best Local Similarity 60.0%; Pred. No. 8.2;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 5 AYDQYGHAAF 14
Db 3 ASDDYGHMRF 12

RESULT 5
LSK2_LEUMA STANDARD; PRT; 10 AA.
ID _LSK2_LEUMA
AC P09039;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Leucosulfakinin-II (LSK-II).
OS Leucophaea maderae (Madeira cockroach), and
OS periplaneta americana (American cockroach).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blaberoidea;
OC Blaberidae; Leucophaea.
OX NCBI_TaxID=6988, 6978;
RN [1]
RP SEQUENCE.
RC SPECIES=L.maderae;
RX MEDLINE=87048769; PubMed=3778455;
RA Nachman R.J., Holman G.M., Cook B.J., Haddon W.F., Ling N.;
RT "Leucosulfakinin-II, a blocked sulfated insect neuropeptide with
RL homology to cholecystokinin and gastrin.";
RL Biochem. Biophys. Res. Commun. 140:357-364(1986).
RN [2]
RP SEQUENCE.
RC SPECIES=P.americana; TISSUE=Corpora cardiaca;
RX MEDLINE=90137190; PubMed=2615921;
RA Veenstra J.A.;
RT "Isolation and structure of two gastrin/CCK-like neuropeptides from
RL the American cockroach homologous to the leucosulfakinins.";
CC -!- FUNCTION: CHANGE THE FREQUENCY AND AMPLITUDE OF CONTRACTIONS OF
CC THE COCKROACH HINDGUT. STIMULATES MUSCLE CONTRACTION OF HINDGUT.
CC -!- SIMILARITY: Belongs to the gastrin/cholecystokinin family.
DR PIR; A26335; GMROL2.
DR PIR; B60656; B60656.
DR InterPro; IPR001651; Gastrin.
DR PROSITE; PS00259; GASTRIN; 1.
KW Hormone; Amidation; Sulfation; Pyrrolidone carboxylic acid.
FT MOD_RES 1 1
FT MOD_RES 5 5
FT MOD_RES 5 5
FT MOD_RES 10 10
FT MOD_RES 10 10
SQ SEQUENCE 10 AA; 1255 MW; 9B4F5391E86B5AAA CRC64;

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:33:34 ; Search time 52 Seconds  
(without alignments)  
86.938 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 459173

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	100.0	16	5	AU98852 E.Coli DN
2	86	100.0	16	6	ABR55126 E. coli d
3	81	94.2	15	2	AAW25795 Peptide d
4	81	94.2	15	3	AAY79543 E. coli d
5	81	94.2	15	4	AAM99341 Vaccine r
6	81	94.2	15	5	AU09842 E.coli dn
7	81	94.2	15	5	AU98876 E. coli D
8	81	94.2	15	5	AAE19457 Human hea
9	81	94.2	15	6	ABR55132 E. coli D
10	81	94.2	16	4	AAM99344 Vaccine r
11	72	83.7	15	2	AAR95446 RA suscep
12	72	83.7	15	2	AAW25796 Peptide d
13	72	83.7	15	5	AAE19458 Heat choc
14	72	83.7	15	6	ABR55133 E. coli d
15	64.5	75.0	14	2	AAR95445 RA suscep
16	54	62.8	15	2	AAW25798 Peptide d
17	54	62.8	15	6	ABR55123 E. coli d
18	34	39.5	9	4	AU03347 Fruit fly
19	34	39.5	9	4	AU03346 Fruit fly
20	34	39.5	9	4	AU03351 Fruit fly
21	34	39.5	9	7	ADE14680 Drosulfak
22	34	39.5	9	7	ADE14676 DmgPCR bi
23	34	39.5	9	7	ADE14679 DmgPCR bi
24	34	39.5	9	7	ADE14677 DmgPCR bi
25	34	39.5	14	3	AAB14081 Human HLA

26	34	39.5	14	4	AAU03353	Aau03353 Fruit fly
27	34	39.5	14	6	AAg79793	Aag79793 Peptide 6
28	34	39.5	14	7	ADE14681	Adel14681 Drosulfak
29	34	39.5	15	2	AAR10086	Aar10086 HLA Dw4 i
30	34	39.5	15	2	AAR10089	Aar10089 HLA Dw4p
31	34	39.5	15	2	AAW25799	Aaw25799 Synthetic
32	34	39.5	15	2	AAW25800	Aaw25800 Synthetic
33	34	39.5	15	3	AAV79544	Aay79544 Human leu
34	34	39.5	15	4	AAW99340	Aam99340 Vaccine r
35	34	39.5	15	5	AAU09843	Aau09843 Human leu
36	34	39.5	15	6	ABR55138	Abrr55138 Human S1
37	34	39.5	15	6	ABR55139	Abrr55139 Human S2
38	34	39.5	15	6	AAg79792	Aag79792 Peptide 6
39	31	36.0	12	3	AAAB40050	Aab40050 Anti-hIL1
40	31	36.0	12	3	AAAB39771	Aab39771 Anti-IL12
41	31	36.0	12	3	AAAB39769	Aab39769 Anti-IL12
42	31	36.0	12	3	AAAB39510	Aab39510 Anti-hIL-
43	31	36.0	14	3	AAAB13800	Aab13800 Human HLA
44	31	36.0	14	3	AAAB14089	Aab14089 Human HLA
45	31	36.0	14	6	AAAg79798	Aag79798 Peptide 6

ALIGNMENTS

RESULT 1	
AAU98852	
ID	AAU98852 standard; peptide; 16 AA.
XX	AAU98852;
AC	
XX	
DT	22-AUG-2002 (first entry)
XX	
DE	E.Coli DNAJ 61 immunogenic peptide.
XX	
KW	Immunogenic peptide; heat shock protein; HSP; DNAJ; immunomodulatory;
KW	cytostatic; antiinflammatory; antibacterial; antiarthritic;
KW	autoimmune disease; arthritis; articular juvenile idiopathic arthritis;
KW	infectious disease; inflammatory bowel disease; cancer;
KW	mucosal tolerisation; DNA vaccination; anergy induction.
XX	
OS	Escherichia coli.
XX	
PN	WO200236611-A2.
XX	
PD	10-MAY-2002.
XX	
PF	31-OCT-2001; 2001WO-US045344.
XX	
PR	01-NOV-2000; 2000US-0245181P.
XX	
PA	(REGC ) UNIV CALIFORNIA.
PA	(MART/) MARTINI A.
XX	
PI	Martini A, Albani S, Carson DA, Prakken BJ;
XX	
DR	WPI; 2002-489999/52.
XX	
PT	New immunomodulatory peptides from heat shock proteins, useful for
PT	treating immunological disorder in subjects such as humans, e.g.
PT	autoimmune disease (e.g. arthritis), infectious disease, inflammatory
PT	bowel disease or cancer.
XX	
PS	Claim 4; Page 55; 84pp; English.
XX	
CC	This invention relates to the use of a peptide, which is an immunogenic
CC	portion derived from a dnaJ heat shock protein (hsp) in modulating an
CC	immune response in a subject. The peptides of the invention may have
CC	immunomodulatory, cytostatic, antiinflammatory, antibacterial or
CC	antiarthritic properties and can stimulate expression of interleukins,
CC	tumour necrosis factor and transforming growth factor beta. The
CC	immunogenic peptide is useful for modulating (i.e. augmenting/inducing or
CC	reducing/inhibiting) an immune response in a subject having an

CC immunological disorder (e.g. autoimmune disease such as arthritis or  
CC articular juvenile idiopathic arthritis), an infectious disease, an  
CC inflammatory bowel disease or cancer. The immunogenic peptide of the  
CC invention is also useful for modulating immunoeffector cell  
CC responsiveness in a subject. The immunogenic peptide is particularly  
CC useful for treating the above-mentioned diseases in mammals, e.g. cat,  
CC dog, horse, farm animal (e.g. ovine, bovine or porcine) or human. In  
CC general, the peptide is useful in methods involving mucosal tolerisation,  
CC DNA vaccination, allergy induction or active immunisation. The present  
CC sequence represents an E. coli DNAJ immunogenic peptide of the invention  
XX  
SQ Sequence 16 AA;

Query Match 100.0%; Score 86; DB 5; Length 16;  
Best Local Similarity 100.0%; Pred. No. 4.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAYDQYGHAAFEQ 16  
DB 1 QKRAYDQYGHAAFEQ 16

RESULT 2  
ABR55126  
ID ABR55126 standard; peptide; 16 AA.

XX ABR55126;

DT 03-JUL-2003 (first entry)

DE E. coli dnaJ61 antigen-specific epitope peptide.

XX Antigen-specific epitope; immune response; T cell; cytokine;  
KW antiarthritic; antidiabetic; neuroprotective; anti-inflammatory;  
KW cyostatic; antithyroid; anti-asthmatic; immunosuppressive;  
KW antipsoriatic; anti-ulcer; antianaemic; cardiant; respiratory;  
KW antiallergic; dermatological; antipsoriatic.

XX Escherichia coli.

XX WO2003026579-A2.

PD 03-APR-2003.

XX 25-SEP-2002; 2002WO-US030578.

XX 25-SEP-2001; 2001US-0325499P.

PR 11-DEC-2001; 2001US-0339284P.

XX (REGC ) UNIV CALIFORNIA.

XX Albani S, Martins A;

XX WPI; 2003-430097/40.

PT Modulating an immune response in a subject having an immune-related  
PT disorder, e.g. arthritis by administering an antigen-specific epitope and  
PT a cytokine or an agent that effects cytokine activity or expression.  
XX  
PS Disclosure; Page 9; 41pp; English.

CC The invention relates to a novel method for modulating an immune response  
CC in a subject having an immune-related disorder. The method comprises: (a)  
CC administering an antigen-specific epitope, where administration provides  
CC epitope-specific T cell immune modulation; and (b) administering a  
CC cytokine, an agent that effects cytokine activity or expression, or an  
CC anticytokine therapy. The method of the invention has antiarthritic,  
CC antidiabetic, neuroprotective, anti-inflammatory, cyostatic,  
CC antithyroid, anti-asthmatic, immunosuppressive, antipsoriatic, anti-  
CC ulcer, antianaemic, cardiant, respiratory general, antiallergic, anti-  
CC dermatological, and antipsoriatic activity. The method is useful for  
CC modulating an immune response in a subject having an immune-related  
CC disorder. The present sequence is used in the exemplification of the

CC invention  
XX  
SQ Sequence 16 AA;

Query Match 100.0%; Score 86; DB 6; Length 16;  
Best Local Similarity 100.0%; Pred. No. 4.1e-08;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAYDQYGHAAFEQ 16  
DB 1 QKRAYDQYGHAAFEQ 16

RESULT 3  
AAW25795  
ID AAW25795 standard; peptide; 15 AA.

XX AAW25795;

DT 06-APR-1998 (first entry)

DE Peptide dnaJ1 which protects against arthritogenic peptides.

XX dnaJ protein; dnaJ1; adult rheumatoid arthritis; vaccine;  
KW arthritogenic peptide; autoimmune disease; systemic immune system;  
KW anti-dnaJ1 antibody; passive immunisation;  
KW rheumatoid arthritis-susceptibility detection.

XX Escherichia coli.

XX WO9734002-A1.

PD 18-SEP-1997.

XX 20-FEB-1997; 97WO-US002957.

XX 15-MAR-1996; 96US-00618464.

XX (REGC ) UNIV CALIFORNIA.

XX Carson DA, Albani S;

XX WPI; 1997-470882/43.

PT Vaccine for protecting against arthritogenic peptide(s) containing dnaJ1  
PT - also determining pre-disposition to rheumatoid arthritis by detecting  
PT anti-arthritogenic peptide antibodies.

XX Disclosure; Page 5; 44pp; English.

CC Peptides AAW25795-96 and AAW25798 are derived from the dnaJ protein from  
CC Escherichia coli. One peptide, dnaJ1 (AAW25795) was found to induce  
CC relatively strong immune response in seropositive adult rheumatoid  
CC arthritis patients. This peptide was used in a vaccine for protecting  
CC against arthritogenic peptides. The vaccine contains a carrier, pure  
CC dnaJ1 peptide or a recombinant gene expression vector encoding the  
CC dnaJ1 peptide. The vaccines are used to prevent rheumatoid arthritis (or  
CC other autoimmune diseases). Vaccines can target the arthritogenic  
CC peptides before they are presented to the systemic immune system. Anti-  
CC dnaJ1 antibodies can also be used for passive immunisation. Antibodies  
CC can be used to detect rheumatoid arthritis-susceptibility sequences on  
CC HLA or in populations of E. coli in the gastrointestinal tract  
XX  
SQ Sequence 15 AA;

Query Match 94.2%; Score 81; DB 2; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.9e-07;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAYDQYGHAAFE 15  
DB 1 QKRAYDQYGHAAFE 15

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:37:19 ; Search time 38 Seconds  
(without alignments)  
132.850 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAVDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 4499

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL 25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	24	27.9	16	10	082407 fragaria vi
2	24	27.9	16	10	082404 fragaria ni
3	24	27.9	16	10	082406 fragaria mo
4	24	27.9	16	10	082402 fragaria nu
5	24	27.9	16	10	082781 fragaria ve
6	24	27.9	16	10	082405 fragaria ii
7	22.5	26.2	16	10	09S8A0
8	22	25.6	15	7	Q9TNP2
9	22	25.6	16	10	Q9S898
10	21	24.4	13	11	Q80Y03
11	21	24.4	16	6	Q9TRR1
12	20	23.3	9	2	Q93LE4
13	20	23.3	11	4	O60614
14	20	23.3	11	5	Q9TWM2
15	20	23.3	11	12	P89269
16	20	23.3	12	6	O46664

17	20	23.3	12	11	Q9QVF2	Q9qvf2 rattus sp.
18	20	23.3	14	10	Q9FYT0	Q9fyto allium cepa
19	20	23.3	15	2	Q9JP39	Q9jp39 pseudomonas
20	20	23.3	15	10	Q9S8I1	Q9s8i1 volvox cart
21	20	23.3	16	10	082403	082403 fragaria ve
22	19.5	22.7	11	2	P96319	P96319 desulfovibr
23	19	22.1	8	6	Q9BF82	Q9bf82 ursus arcto
24	19	22.1	8	6	Q9BFC2	Q9bfc2 macropus eu
25	19	22.1	8	6	Q9BFB9	Q9bfb9 tragelaphus
26	19	22.1	8	6	Q9BFB1	Q9bfb1 echinops te
27	19	22.1	8	6	Q9BFA1	Q9bfai ateles fusc
28	19	22.1	8	6	Q9BFB7	Q9bfb7 tapirus ind
29	19	22.1	8	6	Q9BFB9	Q9bfb9 euphractus
30	19	22.1	8	6	Q9BFB8	Q9bfb8 chaetophrac
31	19	22.1	8	6	Q9BFA8	Q9bfa8 loxodonta a
32	19	22.1	8	6	Q9BFA9	Q9bfa9 procavia ca
33	19	22.1	8	6	Q9BFB2	Q9bfb2 sores arane
34	19	22.1	8	6	Q9BFB5	Q9bfb5 erinaceus c
35	19	22.1	8	6	Q9BFB6	Q9bfb6 myrmecophag
36	19	22.1	8	6	Q9BFB3	Q9bfb3 condylura c
37	19	22.1	8	6	Q9BFB8	Q9bfb8 equus cabal
38	19	22.1	8	6	Q9BFB9	Q9bfb9 roussetus l
39	19	22.1	8	6	Q9BFB4	Q9bfb4 panthera on
40	19	22.1	8	6	Q9BFC3	Q9bfc3 didelphis m
41	19	22.1	8	6	Q9BFA4	Q9bfa4 tupaia mino
42	19	22.1	8	6	Q9BFA2	Q9bfa2 tarsius ban
43	19	22.1	8	6	Q9BFB5	Q9bfb5 leopardus p
44	19	22.1	8	6	Q9BFC1	Q9bfc1 choleopus h
45	19	22.1	8	6	Q9BFB9	Q9bfb9 okapia john

ALIGNMENTS

RESULT 1

082407 PRELIMINARY; PRT; 16 AA.

ID 082407 AC 082407: 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)

DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).

GN ADH.

OS Fragaria viridis (Wild strawberry).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;

OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.

OX NCBI\_TaxID=64942;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=FRA 341;

RA Yu H., Davis T.M.;

RI Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF000220; AAC36547.1; .

DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.

DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-1.; IEA.

DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.

DR GO; GO:0016491; F:oxidoreductase activity; IEA.

KW Oxidoreductase.

FT NON\_TER 1

FT NON\_TER 16

SQ SEQUENCE 16 AA; 1682 MW; CEF73B5B28A4D817 CRC64;

Query Match 27.9%; Score 24; DB 10; Length 16;  
Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
Db 10 YGHEA 14

RESULT 2

082404 ID 082404 PRELIMINARY; PRT; 16 AA.  
AC 082404;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).  
GN ADH.  
OS Fragaria nilgerensis.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
OX NCBI\_TaxID=64941;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berlin 1;  
RA Yu H., Davis T.M.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF000217; AAC36544.1;  
DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-i.; IEA.  
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
KW Oxidoreductase.  
FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1713 MW; CEF73B5B28A4D952 CRC64;  
FT NON\_TER 16  
FT SEQUENCE 16 AA; 1713 MW; CEF73B5B28A4D952 CRC64;  
Query Match 27.9%; Score 24; DB 10; Length 16;  
Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
Db 10 YGHEA 14

RESULT 3  
082406 ID 082406 PRELIMINARY; PRT; 16 AA.  
AC 082406;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).  
GN ADH.  
OS Fragaria moschata.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
OX NCBI\_TaxID=64940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FRA 157;  
RA Yu H., Davis T.M.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF000219; AAC36546.1;  
DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-i.; IEA.  
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
KW Oxidoreductase.  
FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1713 MW; CEF73B5B28A4D952 CRC64;  
Query Match 27.9%; Score 24; DB 10; Length 16;  
Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
Db 10 YGHEA 14

RESULT 4  
082402 ID 082402 PRELIMINARY; PRT; 16 AA.  
AC 082402;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Alcohol dehydrogenase (EC 1.1.1.1) (Fragment).  
GN ADH.  
OS Fragaria rubicola.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
OX NCBI\_TaxID=60188;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FRA520;  
RA Yu H., Davis T.M.;  
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF000213; AAC36540.1;  
DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-i.; IEA.  
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
KW Oxidoreductase.  
FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1682 MW; CEF73B5B28A4D817 CRC64;  
FT NON\_TER 16  
FT SEQUENCE 16 AA; 1682 MW; CEF73B5B28A4D817 CRC64;  
Query Match 27.9%; Score 24; DB 10; Length 16;  
Best Local Similarity 80.0%; Pred. No. 2.9e+03;  
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 YGHAA 13  
Db 10 YGHEA 14

RESULT 5  
082781 ID 082781 PRELIMINARY; PRT; 16 AA.  
AC 082781;  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)  
DT 01-JUN-2003 (TREMBlrel. 24, last annotation update)  
DE Alcohol dehydrogenase (EC 1.1.1.1) (Aldehyde reductase) (Fragment).  
GN ADH.  
OS Fragaria vesca (Woodland strawberry).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids I; Rosales; Rosaceae; Rosoideae; Fragaria.  
OX NCBI\_TaxID=57918;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YELLOW WONDER, and BARON SOLEMACHER;  
RA Yu H., Davis T.M.;  
RL "Genetic relationships among Fragaria species based on RAPDs and an alcohol dehydrogenase (ADH) gene."  
RL Genome 0:0-0(1997).  
CC -1- CATALYTIC ACTIVITY: AN ALCOHOL + NAD(+) = AN ALDEHYDE OR KETONE + NADH.  
CC -1- COFACTOR: ZINC OR IRON.  
CC EMBL; AF000216; AAC36543.1;  
DR EMBL; AF000214; AAC36541.1;  
DR GO; GO:0004025; F:alcohol dehydrogenase activity, iron-dependent; IEA.  
DR GO; GO:0004023; F:alcohol dehydrogenase activity, metal ion-i.; IEA.  
DR GO; GO:0004024; F:alcohol dehydrogenase activity, zinc-dependent; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
KW Oxidoreductase.  
FT NON\_TER 1  
SQ SEQUENCE 16 AA; 1682 MW; CEF73B5B28A4D817 CRC64;  
FT NON\_TER 16  
FT SEQUENCE 16 AA; 1682 MW; CEF73B5B28A4D817 CRC64;



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: July 11, 2004, 09:37:59 ; Search time 23 Seconds  
(without alignments)  
35.914 Million cell updates/sec

Title: US-10-001-938-3  
Perfect score: 86  
Sequence: 1 QKRAAYDQYGHAAFEQ 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 152138

Minimum DB seq length: 0  
Maximum DB seq length: 16

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	94.2	15	1	US-08-618-464-4 Sequence 4, Appli
2	81	94.2	15	3	US-09-107-615-4 Sequence 4, Appli
3	81	94.2	15	5	PCT-US95-04896-4 Sequence 4, Appli
4	72	83.7	15	3	US-08-618-464-5 Sequence 5, Appli
5	72	83.7	15	3	US-09-107-615-5 Sequence 5, Appli
6	72	83.7	15	5	PCT-US95-04896-5 Sequence 5, Appli
7	54	62.8	15	1	US-08-618-464-6 Sequence 6, Appli
8	54	62.8	15	3	US-09-107-615-6 Sequence 6, Appli
9	34	39.5	15	1	US-08-618-464-7 Sequence 7, Appli
10	34	39.5	15	1	US-08-618-464-8 Sequence 8, Appli
11	34	39.5	15	3	US-09-107-615-7 Sequence 7, Appli
12	34	39.5	15	3	US-09-107-615-8 Sequence 8, Appli
13	31	36.0	12	1	US-08-170-114A-5 Sequence 5, Appli
14	31	36.0	12	1	US-08-170-114A-5 Sequence 5, Appli
15	31	36.0	14	2	US-08-323-686-3 Sequence 3, Appli
16	31	36.0	14	2	US-08-323-686-26 Sequence 26, Appli
17	31	36.0	15	4	US-09-255-501-67 Sequence 67, Appli
18	31	36.0	15	4	US-09-255-501-68 Sequence 68, Appli
19	30	34.9	9	4	US-09-721-870-180 Sequence 180, App
20	30	34.9	15	3	US-08-467-023-255 Sequence 255, App
21	28	32.6	11	3	US-08-467-580-14 Sequence 14, Appli
22	28	32.6	11	5	PCT-US95-08516-14 Sequence 14, Appli
23	28	32.6	15	2	US-08-521-871A-3 Sequence 3, Appli
24	27	31.4	12	1	US-07-732-114A-11 Sequence 11, Appli
25	27	31.4	12	1	US-08-170-114A-11 Sequence 11, Appli
26	27	31.4	13	1	US-08-487-568-45 Sequence 45, Appli
27	26	30.2	6	1	US-07-995-503A-12 Sequence 12, Appli

28	26	30.2	6	1	US-08-390-510-12 Sequence 12, Appli
29	26	30.2	6	1	US-08-390-790-12 Sequence 12, Appli
30	26	30.2	6	2	US-08-390-509-12 Sequence 12, Appli
31	26	30.2	6	3	US-09-149-860A-12 Sequence 12, Appli
32	26	30.2	12	1	US-07-995-503A-11 Sequence 11, Appli
33	26	30.2	12	1	US-08-390-510-11 Sequence 11, Appli
34	26	30.2	12	1	US-08-390-790-11 Sequence 11, Appli
35	26	30.2	12	2	US-08-390-509-11 Sequence 11, Appli
36	26	30.2	12	3	US-09-149-860A-11 Sequence 11, Appli
37	26	30.2	13	2	US-08-194-981E-8 Sequence 11, Appli
38	26	30.2	15	2	US-08-521-871A-4 Sequence 8, Appli
39	26	30.2	15	3	US-09-613-182-1 Sequence 1, Appli
40	26	30.2	15	4	US-09-490-702B-19 Sequence 19, Appli
41	26	30.2	15	4	US-09-490-702B-96 Sequence 96, Appli
42	26	30.2	15	4	US-09-255-501-69 Sequence 69, Appli
43	26	30.2	16	3	US-09-181-896-2 Sequence 2, Appli
44	26	30.2	16	4	US-09-490-702B-95 Sequence 95, Appli
45	25	29.1	9	1	US-08-186-266-11 Sequence 11, Appli

ALIGNMENTS

RESULT 1  
US-08-618-464-4  
Sequence 4, Application US/08618464  
Patent No. 5773570  
GENERAL INFORMATION:  
APPLICANT: CARSON, DENNIS A.  
APPLICANT: ALBANI, SALVATORE  
TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN IMMUNE PROTECTION AGAINST ARTHRITIC  
TITLE OF INVENTION: INDUCING  
TITLE OF INVENTION: INVOLVED IN  
TITLE OF INVENTION: THE PATHOGENESIS OF RHEUMATOID ARTHRITIS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & RICHARDSON P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CALIFORNIA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/618,464  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: HOWELLS, STACY L.  
REGISTRATION NUMBER: 34,842  
REFERENCE/DOCKET NUMBER: 07340/042001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE:  
CLONE: Immunogenic dna/ Peptide  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..15  
US-08-618-464-4  
Query Match 94.2%; Score 81; DB 1; Length 15;

Best Local Similarity 100.0%; Pred. No. 5.8e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 QKRAAYDQYGHAAFE 15  
Db 1 QKRAAYDQYGHAAFE 15

## RESULT 2

US-09-107-615-4  
; Sequence 4, Application US/09107615  
; Patent No. 6153200

## GENERAL INFORMATION:

APPLICANT: CARSON, DENNIS A.  
APPLICANT: ALBANI, SALVATORE  
TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN  
INDUCING IMMUNE PROTECTION AGAINST  
TITLE OF INVENTION: ARTHRITIS  
TITLE OF INVENTION: THE PATHOGENESIS OF RHEUMATOID ARTHRITIS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & RICHARDSON P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CALIFORNIA  
COUNTRY: US  
ZIP: 92037

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/107,615  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/618,464  
FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: HOWELLS, STACY L.  
REGISTRATION/DOCKET NUMBER: 34,842  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619/678-5070  
TELEFAX: 619/678-5099

## INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

IMMEDIATE SOURCE: Immunogenic dnaJ Peptide

FEATURE: NAME/KEY: Peptide

LOCATION: 1..15

US-09-107-615-4

Query Match 94.2%; Score 81; DB 3; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.8e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFE 15

Db 1 QKRAAYDQYGHAAFE 15

## RESULT 3

PCT-US95-04896-4

; Sequence 4, Application PC/TUS9504896  
; GENERAL INFORMATION:

APPLICANT: THE REGENTS OF THE UNIVERSITY  
APPLICANT: OF CALIFORNIA  
TITLE OF INVENTION: METHOD AND REAGENTS FOR THE TREATMENT  
TITLE OF INVENTION: OF RHEUMATOID ARTHRITIS  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ROBBINS, BERLINER & CARSON  
STREET: 201 N. FIGUEROA STREET, 5TH FLOOR  
CITY: LOS ANGELES  
STATE: CALIFORNIA  
COUNTRY: US  
ZIP: 90012  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/04896  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: BERLINER, ROBERT  
REGISTRATION/DOCKET NUMBER: 20,121  
REFERENCE/DOCKET NUMBER: 5555-314  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 213-977-1001  
TELEFAX: 213-977-1003  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
IMMEDIATE SOURCE: Immunogenic dnaJ Peptide  
CLONE: Immunogenic dnaJ Peptide  
FEATURE: NAME/KEY: Peptide  
LOCATION: 1..15  
PCT-US95-04896-4

Query Match 94.2%; Score 81; DB 5; Length 15;  
Best Local Similarity 100.0%; Pred. No. 5.8e-08;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QKRAAYDQYGHAAFE 15  
Db 1 QKRAAYDQYGHAAFE 15

RESULT 4  
US-08-618-464-5  
; Sequence 5, Application US/08618464  
; Patent No. 5773570  
; GENERAL INFORMATION:  
APPLICANT: CARSON, DENNIS A.  
APPLICANT: ALBANI, SALVATORE  
TITLE OF INVENTION: VACCINE COMPOSITIONS AND METHODS USEFUL IN  
INDUCING IMMUNE PROTECTION AGAINST ARTHRITIS  
TITLE OF INVENTION: INVOLVED IN  
TITLE OF INVENTION: THE PATHOGENESIS OF RHEUMATOID ARTHRITIS  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & RICHARDSON P.C.  
STREET: 4225 Executive Square, Suite 1400  
CITY: La Jolla  
STATE: CALIFORNIA  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible